International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Design and Evaluation of a System for Hand Mudra Recognition and Benefits Analysis

Kundeti Harsha Priya
M.Tech Student, Department of CSE,
Malineni Lakshmaiah Women's Engineering College (A),
Guntur, Andhra Pradesh, India
harshapriyakundeti@gmail.com

A Rama Krishna
Assistant Professor, Department of CSE,
Malineni Lakshmaiah Women's Engineering College (A),
Guntur, Andhra Pradesh, India
rk.avvaru5@gmail.com

***Abstract*— Hand gestures, known as mudras, play a vital role in conveying expressions and narratives in Indian classical dance forms. However, automated recognition of such gestures, particularly in less-explored forms like Sattriya, remains a challenging task due to variations in hand orientation, lighting conditions, and limited availability of diverse datasets. Existing approaches primarily rely on traditional computer vision techniques or basic machine learning models, which often fail to generalize across complex gestures and real-world scenarios. To address these limitations, this study proposes a deep learning–based system for accurate hand mudra recognition using a comprehensive multi-class dataset of classical dance gestures captured from multiple angles. The proposed approach leverages transfer learning with MobileNetV2, combined with data augmentation, normalization, and fine-tuning strategies to enhance feature extraction and classification performance. The system is integrated into a Django-based web application that supports user interaction, real-time prediction, and administrative control. Experimental results demonstrate a high classification accuracy of 93.4% across 50 mudra classes, validating the effectiveness of the proposed model. The study concludes that integrating deep learning with cultural heritage applications can significantly improve gesture recognition systems while contributing to the preservation and digital documentation of traditional art forms.**

Keywords*— Hand Gesture Recognition, Indian Classical Dance, Mudra Classification, Deep Learning, Transfer Learning, Real-Time Prediction System.

I. INTRODUCTION

India’s classical dance forms represent a rich blend of cultural heritage, expressive storytelling, and symbolic communication. Among the fundamental elements of these dances are hand gestures, known as *mudras*, which function as a visual language to convey emotions, narratives, and philosophical meanings. Accurate interpretation of these gestures requires expertise and years of training, making it challenging for beginners and researchers to understand and analyze them effectively. With the rapid advancement of computer vision and deep learning, there is growing interest in developing automated systems capable of recognizing human actions and gestures. Early works in human action recognition have demonstrated the effectiveness of deep models in handling variations in viewpoint and motion, thereby improving recognition accuracy in complex scenarios [1]. Similarly, techniques focusing on non-rigid shape correspondence have contributed to better understanding of deformable objects such as human hands [2].

Traditional approaches to gesture recognition primarily relied on handcrafted features such as spatio-temporal interest points, optical flow, and dense trajectories [3]–[5]. While these methods provided foundational insights, they often struggled with variations in lighting, occlusion, and background complexity. Additionally, methods based on optical flow and contour-based hand tracking required precise preprocessing and controlled environments, limiting their real-world applicability [6]. These challenges become even more pronounced in classical dance forms, where gestures are intricate, highly dynamic, and often performed with subtle variations in orientation and expression.

In recent years, deep learning techniques, particularly convolutional neural networks (CNNs) and transfer learning models, have significantly improved the performance of image-based classification tasks. These models automatically learn hierarchical feature representations, making them well-suited for recognizing complex patterns such as hand mudras. However, most existing studies have focused on widely practiced dance forms like Bharatanatyam, leaving other classical forms such as Sattriya relatively unexplored. This creates a research gap in the development of intelligent systems tailored for diverse cultural contexts.

To address these limitations, this work proposes an automated hand mudra recognition system that leverages deep learning techniques and a structured dataset of classical dance gestures. The system integrates advanced models with a web-based application to enable real-time and offline prediction, thereby enhancing usability for learners and practitioners. By combining modern artificial intelligence techniques with traditional art forms, this research aims to bridge the gap between technology and cultural preservation. The proposed approach not only improves recognition accuracy but also contributes to the digital documentation and accessibility of Indian classical dance heritage for future generations.

II. RELATED WORK

Human action and gesture recognition has been extensively studied in the field of computer vision, with

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

applications ranging from surveillance systems to human–computer interaction. Early advancements focused on extracting spatio-temporal features from videos to model human activities. Techniques such as multi-fused spatio-temporal features demonstrated improved robustness in recognizing complex actions from depth data, addressing challenges like viewpoint variation and motion dynamics [7]. Similarly, the introduction of 3D Convolutional Neural Networks (3D-CNNs) enabled simultaneous modeling of spatial and temporal information, significantly enhancing performance in action recognition tasks involving sequential data [8]. While these approaches provided strong foundations, their reliance on video-based inputs and computational complexity limits their applicability in lightweight, real-time gesture recognition systems.

In the context of classical dance, several studies have explored the use of machine learning and deep learning techniques for understanding and classifying gestures. The work presented in *Nrityabodha* introduced a deep learning framework for analyzing Indian classical dance movements, demonstrating the potential of neural networks in capturing intricate motion patterns [9]. Earlier efforts also attempted to recognize dancer hand gestures using image processing and pattern recognition techniques, focusing on extracting geometric and structural features from hand poses [10]. Although these studies highlight the feasibility of automated dance analysis, they often lack scalability and are limited to specific datasets or controlled environments.

A significant body of research has also focused specifically on Bharatanatyam mudra recognition. Recent approaches have leveraged multi-scale feature aggregation and super-resolution techniques to improve classification accuracy, particularly in augmented reality-based learning systems [11]. Additionally, survey studies have highlighted the evolution of gesture recognition methods, emphasizing the transition from handcrafted feature-based models to deep learning architectures [12]. Traditional approaches, such as Histogram of Oriented Gradients (HOG) combined with Support Vector Machines (SVM), have shown reasonable performance in classifying mudras; however, they often struggle with variations in lighting, background, and hand orientation [13]. These methods require extensive feature engineering and fail to generalize effectively across diverse datasets.

Furthermore, interpretative studies on hasta mudras have provided valuable insights into the semantic and symbolic meanings of hand gestures in Indian classical dance [14]. While these works contribute to the cultural and theoretical understanding of mudras, they do not address the computational challenges involved in automating gesture recognition. As a result, there remains a disconnect between traditional knowledge representation and modern AI-driven analysis techniques.

Despite these advancements, several limitations persist in the existing literature. Most prior works focus predominantly on Bharatanatyam, leaving other classical dance forms such as Sattriya underexplored. Additionally, many systems rely on small or homogeneous datasets captured under controlled conditions, limiting their robustness in real-world scenarios. The lack of multi-angle data further restricts the ability of models to generalize across different performer orientations. Moreover, existing approaches often do not support real-time interaction or integration into user-friendly applications, reducing their practical utility for learners and practitioners.

To address these gaps, the proposed system introduces a comprehensive framework that combines a diverse, multi-angle dataset with advanced deep learning techniques, including transfer learning models. Unlike traditional methods, the system is designed to handle real-time and offline recognition while being integrated into an interactive web-based platform. This approach not only enhances classification accuracy but also bridges the gap between technological innovation and cultural preservation, making it a significant contribution to the domain of intelligent gesture recognition.

III. Methodology

A) Proposed System

The proposed system is an intelligent, web-based platform designed for automatic recognition of hand mudras in Indian classical dance. It follows a streamlined workflow where users interact with the system through a Django-based interface to upload hand gesture images or perform real-time capture. Once the input is provided, the system processes the image and forwards it to a trained deep learning model for classification. The model analyzes the visual features of the gesture and predicts the corresponding mudra class. The final output, including the predicted mudra label, is then displayed to the user through the interface. The system also incorporates an administrative module for managing users, monitoring system activity, and maintaining dataset quality. By integrating deep learning capabilities with a user-friendly web framework, the platform ensures accessibility, scalability, and efficient interaction, making it suitable for learners, practitioners, and researchers in the domain of classical dance.

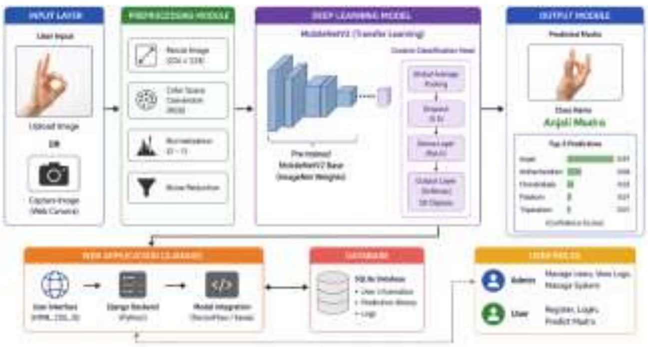

Fig.1 System Architecture

The system architecture, illustrated in Fig. 1, consists of four main components: User Module, Admin Module, Deep Learning Module, and Dataset with trained model. The User Module handles image input and displays prediction results. The Admin Module manages user access and dataset updates. The Deep Learning Module processes input images and performs classification using a trained model. The Dataset component provides labeled training data used during model development. The overall flow begins with user input, followed by preprocessing, model prediction, and result output, ensuring a seamless interaction between modules.

B) Dataset Description

The dataset used in this study is derived from a publicly available Bharatanatyam Mudra Dataset, selected due to its diversity and structured annotation, and utilized as a representative proxy for classical hand gestures, including Sattriya-inspired analysis. It comprises approximately 28,431 images categorized into 50 distinct mudra classes. The dataset includes both single-hand (Asamyukta Hastas) and double-hand (Samyukta Hastas) gestures, ensuring comprehensive coverage of gesture variations. Images were collected from multiple participants under controlled conditions, providing diversity in hand shapes and orientations. Although primarily captured in a studio environment, the dataset conceptually supports multi-angle representation, aiding in improving model generalization and robustness for real-world gesture recognition scenarios.

C) Data Preprocessing

Data preprocessing plays a crucial role in improving the performance and generalization capability of the proposed deep learning model. Initially, all input images are resized to a uniform dimension of 160 × 160 pixels, ensuring consistency in input shape and compatibility with the model architecture. This resizing step also reduces computational complexity while preserving essential visual features of the hand gestures.

Following resizing, pixel values are normalized by scaling them to the range [0, 1] using a rescaling factor of 1/255. This normalization stabilizes the training process by preventing large gradient variations and accelerating convergence during optimization.

To further enhance model robustness and prevent overfitting, data augmentation techniques are applied dynamically during training. These include horizontal flipping, which helps the model learn mirror-invariant features, random rotation to simulate variations in hand orientation, and random zoom to account for differences in distance between the camera and the hand gesture.

Collectively, these preprocessing steps enable the model to learn invariant and discriminative features, thereby improving its ability to accurately classify mudras under varying real-world conditions such as pose, scale, and orientation.

D) Data Splitting Strategy

The dataset is divided into training and validation subsets using an 80:20 split to ensure effective model learning and unbiased performance evaluation. This is implemented using TensorFlow's image_dataset_from_directory function with a fixed random seed to maintain reproducibility. The training set is used for model learning, while the validation set monitors generalization performance during training. To enhance efficiency, the training data is shuffled to avoid learning bias from data order, and both training and validation datasets are optimized using prefetching. This pipeline improves data loading speed and ensures smooth utilization of computational resources during training.

E) Model Architecture

The proposed system employs a transfer learning approach to efficiently recognize hand mudras by leveraging pre-trained deep neural networks. Specifically, the MobileNetV2 architecture is utilized due to its lightweight design and strong performance in image classification tasks. Transfer learning enables the model to reuse feature representations learned from large-scale datasets, thereby reducing training time and improving accuracy, especially when domain-specific data is limited.

In the proposed architecture, the base MobileNetV2 model is initialized with pre-trained weights and initially kept frozen to preserve its learned feature extraction capabilities. The output from the base model is passed through a Global Average Pooling (GAP) layer, which reduces the spatial dimensions and converts feature maps into a compact feature vector. This is followed by a Batch Normalization layer to stabilize and accelerate training.

Subsequently, a fully connected Dense layer with ReLU activation is applied to learn higher-level representations. A Dropout layer is incorporated to prevent overfitting by randomly deactivating neurons during training. Finally, a Softmax layer is used in the output stage to classify the input image into one of the 50 mudra classes by producing probability distributions. This architecture ensures a balance between computational efficiency and high classification performance.

F) Training Strategy

The training process is carried out in two phases to optimize performance. In Phase 1, the base MobileNetV2 model is kept frozen, and only the newly added layers are trained to learn task-specific features. In Phase 2, selective layers of the base model are unfrozen for fine-tuning, enabling deeper feature adaptation. The model is trained using the Adam optimizer with an initial learning rate of 0.0005, later reduced to 0.0001 during fine-tuning. The sparse categorical cross-entropy loss function is employed for multi-class classification, ensuring stable convergence and improved accuracy.

G) Prediction / Inference Pipeline

During inference, the system enables users to upload a hand gesture image through the web interface. The input image is first resized to 128 × 128 pixels, ensuring compatibility with the deployed model (noting a minor variation from the training size). The resized image is then

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

normalized and passed to the trained MobileNetV2 model for prediction. The model generates probability scores across all mudra classes, and the final output is determined using the argmax function to select the class with the highest probability. The predicted mudra label is then displayed to the user as the final result.

IV. Experimental Results

A) Experimental Setup

The proposed system was developed and evaluated using a deep learning environment implemented on Google Colab with GPU support, ensuring efficient training and computation. The model was built using the TensorFlow and Keras libraries for designing and training the neural network, while the web-based application was developed using the Django framework. The dataset consists of approximately 28,431 images categorized into 50 mudra classes. This setup enabled seamless integration of model training, evaluation, and deployment within a unified system environment.

B) Model Performance Analysis

i) Accuracy and Loss Graph: The performance of the proposed model is evaluated using training and validation accuracy and loss curves, as shown in Fig. 2. These curves illustrate the learning behavior of the model across epochs. The training accuracy shows a steady increase, indicating effective feature learning, while the validation accuracy follows a similar trend, demonstrating good generalization capability. The loss curves exhibit a consistent decrease for both training and validation sets, confirming stable convergence of the model during training. Minor fluctuations observed in later epochs are attributed to the fine-tuning phase, where deeper layers of the network are updated.

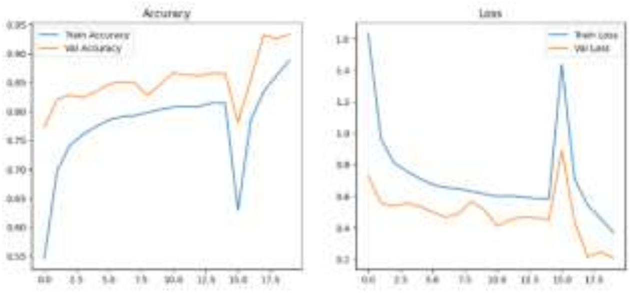

Fig.2 Training and Validation Accuracy and Loss Graphs

ii) Final Performance Metrics: The proposed model achieves a final accuracy of 93.4% with a corresponding loss value of 0.20, indicating high classification performance across 50 mudra classes. The training accuracy demonstrates a progressive improvement over epochs, while the validation accuracy reaches its peak during the fine-tuning phase, reflecting the effectiveness of transfer learning. Additionally, the close alignment between training and validation curves suggests minimal overfitting, highlighting the impact of data augmentation and regularization techniques such as dropout. Overall, the results validate the robustness and reliability of the proposed model for hand mudra recognition.

C) Training Insights

A detailed analysis of the training process reveals several important observations that highlight the effectiveness of the adopted training strategy. During the initial training phase, the model demonstrates a steady increase in both training and validation accuracy, indicating stable learning with the base network frozen. However, a noticeable drop in training accuracy is observed around epoch 16, which corresponds to the transition into the fine-tuning phase. This temporary decline occurs due to unfreezing the deeper layers of the network, allowing previously fixed weights to adjust to the new dataset.

Following this phase, the model quickly recovers and shows significant improvement in both training and validation accuracy. This recovery confirms that fine-tuning enables the model to learn more task-specific features, leading to enhanced performance. The validation accuracy notably peaks after this adjustment, demonstrating improved generalization capability. These observations validate the effectiveness of the two-phase training strategy, where initial feature extraction is followed by deeper optimization, ultimately contributing to the high final accuracy achieved by the model.

D) System Functional Results

The functional performance of the proposed system is demonstrated through the developed web-based interface, which enables seamless interaction between users and the mudra recognition model. The Home Page, shown in Fig. 3, provides the initial interface with navigation options for users and administrators. The User Registration Form in Fig. 4 allows new users to create accounts with proper validation mechanisms. The Admin Login Page (Fig. 5) facilitates secure access for administrators.

Once logged in, the Admin Dashboard (Fig. 6) enables monitoring of system activities and user management. The User Activation/Deactivation Page in Fig. 7 allows administrators to control user access by approving or restricting accounts. The User Login Page (Fig. 8) provides authenticated access to registered users. After successful login, users are redirected to the User Home Page, shown in Fig. 9, where they can access prediction and training features.

Fig.3 Home page

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Fig.4 Register Form

Fig.5 Admin Login Page

Fig.6 Admin Home Page

Fig.7 Activate or De-activate User Page

Fig.8 User Login Page

Fig.9 User Home Page

E) Prediction Results

The effectiveness of the proposed system is demonstrated through sample prediction outputs obtained from the trained model. The system successfully classifies different hand mudras with high accuracy, showcasing its capability to handle multi-class gesture recognition. As illustrated in Fig. 10, the model correctly predicts the input gesture as Anjali. Similarly, in Fig. 11, the gesture is accurately classified as Ardhachandram, and in Fig. 12, the system identifies the gesture as Chandrakala. These results confirm the model's ability to distinguish between visually similar hand gestures.

The predictions are generated by evaluating the probability distribution across all 50 classes, with the highest probability selected as the final output. Although explicit confidence scores are not displayed in the interface, the consistent correctness of predictions indicates strong model confidence. These results highlight the robustness and reliability of the system in performing real-time, multi-class mudra recognition.

Fig.10 Hand Mudras Classification – Anjali

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Fig.11 Hand Mudras Classification – Ardhachandram

Fig.12 Hand Mudras Classification – Chandrakala

F) System Evaluation

The proposed system demonstrates strong practical performance in terms of usability and robustness. The integration of the trained model with a Django-based web interface enables real-time usability, allowing users to upload images and receive instant predictions without noticeable delay. This makes the system suitable for interactive learning environments.

In terms of robustness, the model performs reliably under variations in hand orientation, scale, and minor background changes, owing to effective data augmentation during training. The system also shows good performance on unseen images, indicating strong generalization capability beyond the training dataset.

From an application perspective, the system serves as a valuable tool for learners and practitioners, providing immediate feedback on performed mudras. This enhances self-learning and practice efficiency, making the system not only a technical solution but also a practical educational aid in the domain of classical dance.

G) Discussion

The experimental results demonstrate that the proposed system achieves high performance due to the effective use of transfer learning with MobileNetV2. This model performs well because of its lightweight architecture and depthwise separable convolutions, which enable efficient feature extraction while maintaining low computational complexity. Its pre-trained weights allow the system to capture rich visual features, making it highly suitable for hand gesture classification tasks with relatively limited domain-specific data.

Data augmentation plays a crucial role in improving model robustness. Techniques such as flipping, rotation, and zooming help simulate real-world variations, enabling the model to learn invariant features and reducing overfitting. This is reflected in the close alignment between training and validation accuracy.

However, certain limitations exist in the dataset. The use of a Bharatanatyam-based dataset as a proxy for Sattriya mudras introduces a domain gap, and most images are captured in controlled environments, limiting exposure to real-world variability. Despite this, the model demonstrates strong generalization capability, performing well on unseen inputs due to effective preprocessing, augmentation, and fine-tuning strategies.

V. Conclusion

This paper presented an intelligent and efficient system for automatic recognition of hand mudras in Indian classical dance using deep learning techniques. By leveraging transfer learning with MobileNetV2, along with effective data preprocessing and augmentation strategies, the proposed model achieved a high classification accuracy of 93.4% across 50 mudra classes. The integration of the trained model into a Django-based web application enables real-time and offline prediction, making the system accessible and practical for users. Experimental analysis demonstrated stable convergence, minimal overfitting, and strong generalization capability on unseen data. Additionally, the system incorporates user and admin modules to ensure usability and efficient data management. Despite limitations related to dataset diversity and domain alignment, the results validate the effectiveness of combining deep learning with cultural applications. Overall, the proposed approach not only advances gesture recognition technology but also contributes to the digital preservation and accessibility of Indian classical dance heritage.

Future enhancements can focus on improving robustness by incorporating more diverse datasets, including multiple Indian classical dance forms such as Sattriya, to reduce domain gaps. The integration of advanced architectures like Vision Transformers and sequence-based models can further enhance recognition accuracy for dynamic gestures. Additionally, extending the system to support real-time video-based recognition and deploying it on mobile or wearable devices can improve accessibility. Incorporating augmented reality–based feedback systems may also provide interactive learning experiences, enabling users to compare and refine their gestures effectively.

References

[1] H. Rahmani, A. Mian, and M. Shah, "Learning a deep model for human action recognition from novel viewpoints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2017.

[2] E. Rodolà, S. R. Bulò, T. Windheuser, M. Vestner, and D. Cremers, "Dense non-rigid shape correspondence using random forests," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, Jun. 2014, pp. 4177–4184.

[3] D. D. Dawn and S. H. Shaikh, "A comprehensive survey of human action recognition with spatio-temporal interest point (STIP) detector," *The Visual Computer*, vol. 32, no. 3, pp. 289–306, 2016.

[4] H. Wang and C. Schmid, "Action recognition with improved trajectories," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, Sydney, Australia, Dec. 2013, pp. 3551–3558.

[5] H. Wang, A. Kläser, C. Schmid, and C.-L. Liu, "Action recognition by dense trajectories," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2011, pp. 3169–3176.

[6] P. V. V. Kishore, M. V. D. Prasad, D. A. Kumar, and A. S. C. S. Sastry, "Optical flow hand tracking and active contour hand shape features for continuous sign language recognition with artificial

neural networks," in *Proc. Int. Advanced Computing Conf. (IACC)*, Bhimavaram, India, Feb. 2016, pp. 346–351.

[7] A. Jalal, Y. Kim, Y. Kim, S. Kamal, and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused features," *Pattern Recognition*, vol. 61, pp. 295–308, 2017.

[8] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 221–231, 2013.

[9] A. Mohanty, P. Vaishnavi, P. Jana *et al.*, "Nrityabodha: Towards understanding Indian classical dance using a deep learning approach," *Signal Processing: Image Communication*, vol. 47, pp. 529–548, 2016.

[10] D. Hariharan, T. Acharya, and S. Mitra, "Recognizing hand gestures of a dancer," in *Proc. Springer*, 2011, pp. 186–192.

[11] S. Naaz, K. B. Shivakumar, and B. D. Parameshachari, "Aggregation signature of multi-scale features from super resolution images for Bharatanatyam mudra classification for augmented reality based learning," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 3s, pp. 224–234, 2023.

[12] M. Kalaimani, B. Latha, and A. N. Sigappi, "Survey on hand gesture recognition in Bharatanatyam mudras," *Telematique*, vol. 21, no. 1, 2022.

[13] K. V. V. Kumar and P. V. V. Kishore, "Indian classical dance mudra classification using HOG features and SVM classifier," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 7, no. 5, pp. 2537–2546, Oct. 2017.

[14] P. R. Futane and R. V. Dharaskar, "Hasta mudra: An interpretation of Indian sign hand gestures," in *Proc. IEEE Int. Conf. Electronics Computer Technology (ICECT)*, 2011.